OSU Working Papers in Linguistics 44 (70-91)

The Relationship between Syntactic and Semantic Processes in Sentence Comprehension*

Julie E. Boland
jboland@ling.ohio-state.edu

Abstract: Two experiments investigate how lexically ambiguous input is handled by the sentence processing system and shed light on the relationship between syntactic and semantic processing. Sentence contexts containing ambiguous verbs (e.g., *Which salad/baseball did Janet toss...* probe-word: *Bill*) are used to investigate how subcategorization and thematic role information is used by the sentence processing system. The results are consistent with a model in which the syntactic processing system uses subcategorization information to compute all "legal" structures in parallel, without consideration of semantic information from the context. Meanwhile, the semantic processing system uses contextual information to pursue the single most likely semantic analysis. The resulting syntactic and semantic representations are checked against each other, and inconsistent analyses discarded.

Models of sentence understanding often decompose the task into syntactic and semantic processes. The degree to which syntactic and semantic processes are independent, and the relationship between them, are matters of great debate. Tangled up in this debate is the role of combinatory lexical information in the initial stages of sentence understanding. This is because a verb's "subcategorization frames" and "thematic structure" are part of the lexical knowledge made available when a verb is recognized. A subcategorization frame is a syntactic representation of a verb's arguments.[1] A thematic structure is a representation of the "thematic roles" assigned by a verb, thematic roles being generalized characterizations of an argument's mode of participation in the event described by a sentence.

Views regarding how combinatory lexical information is used by the sentence processing system vary widely. The "Garden Path" model maintains that there is a

*Experiment 1 was completed at the University of Rochester as part of the author's doctoral dissertation and was supported by grant HD-27206 (to Michael Tanenhaus). Experiment 2 was conducted at the Ohio State University and was supported by a University Seed Grant. The author gratefully thanks Michael Tanenhaus, Chris Barker, Kim Darnell, and three anonymous reviewers for many helpful comments on early versions of this manuscript.

[1] I am distinguishing arguments from adjuncts, which are always optional, and not subcategorized.

syntactic "module" which operates independently of semantic knowledge, excluding contextual and most lexical information from initial syntactic processing (e.g., Frazier, 1987, 1989, Mitchell, 1989). The initial parse is influenced only by major syntactic category, phrase structure rules, and a heuristic specifying that the simplest structure will always be constructed first. Because relevant information is initially ignored, the parser will make frequent mistakes, or "wander down the garden path". The opposing view is that the syntactic and semantic systems are completely integrated, with the extreme version making no distinction between syntactic and semantic processes (Bates & MacWhinney, 1982; Waltz & Pollack, 1985). Marslen-Wilson and Tyler (1987) take a more moderate position, in which theoretically distinct syntactic and semantic systems work interactively. A similar position has been taken by Tanenhaus and colleagues (e.g., Tanenhaus et al., 1989). They argue that subcategorization and thematic information are both used in the early stages of sentence comprehension, leading to the prediction that garden paths should occur only when lexical knowledge is ambiguous.

The current paper presents a processing model that requires still more moderation of the Interactive position. In the proposed model, the syntactic processing system functions independently of semantic information from the context. However, unlike the Garden Path model, the syntactic processing system uses subcategorization information to compute all "legal" structures in parallel. Meanwhile, the semantic processing system uses information from the output of the syntactic system and a variety of other sources (the lexicon, the discourse model, etc) to pursue the most likely semantic analysis.[2] The proposed model might be termed a "Concurrent" model, because the syntactic and semantic representations are computed simultaneously.

The Concurrent model is appealing on theoretical grounds for several reasons. First, because context is ignored by the syntactic system, syntactic representations can be built reflexively using phrase structure rules (or the equivalent) and syntactic category and subcategory information from the lexicon. When multiple structures are consistent with lexical information, multiple structures are automatically built. Little cost should result from building multiple representations because the structures are output automatically, in parallel. In contrast, semantic representations are heavily context-dependent, and their construction is a resource-intensive process. However, precisely *because* they are context-dependent, the context guides the construction of the single most likely semantic representation. If the initial semantic representation proves inconsistent with new information (such as the following words), the alternative syntactic representations may provide a mechanism for recovery.

The Concurrent model is similar to a class of models that can be described as "Parallel Syntax" models. A growing number of researchers are finding evidence for parallel postulation of syntactic structures, possibly using subcategorization

[2]In future versions of this model, it may be possible to specify how various properties of the context bear on the decision of which semantic analysis to pursue, but currently, only "thematic" properties are considered. When the verb is recognized, the set of possible thematic frames are compared with the current syntactic representation and the semantic properties of the current arguments. The best-fitting thematic frame is selected as the basis for the semantic analysis.

information, just like the Concurrent model (e.g., Crain & Steedman, 1985; Gorrell, 1989 & 1991; McElree, 1993). However, most advocates of the parallel models seem to assume, as does Frazier (1987), that the syntactic analysis must precede the semantic analysis, and some of them make specific claims about the order in which the "parallel" syntactic representations undergo semantic processing. Like the Garden Path model, these models are essentially modular. Under the Concurrent model, it is possible for a semantic representation to be produced before the parser has settled on a single syntactic structure. Such an outcome is likely when subcategorization information permits computation of multiple syntactic representations, but contextual information strongly biases a single interpretation. While such an outcome *may* be possible with a Parallel Syntax model, it is clearly not possible with the Garden Path model. The current experiments do not allow comparisons that would clearly distinguish between the Concurrent model and some Parallel Syntax models. Therefore, the contrast will be between the Concurrent model, the Garden Path model, and Interactive models in which semantic information influences the initial parse.

In order to distinguish between the Concurrent model, the Garden Path model, and Interactive models, one must have the methodological tools to distinguish between syntactic and semantic representations in a way that is not confounded with theoretical assumptions. For example, it has sometimes been implied that first pass eye-movements are indicative of first-pass (i.e., syntactic) parsing and regressive/second pass eye-movements are indicative of later (i.e., semantic) interpretive processes (e.g., Ferreira & Henderson, 1990). While this is perfectly plausible, it is linked to the theoretical claim that syntactic processing precedes semantic processing. There are no pre-theoretical grounds for identifying first pass eye movements with purely syntactic processes. In contrast, recent work measuring Event-Related Brain Potentials (ERPs) has suggested that distinct components of the waveform can be linked to syntactic and semantic anomalies (Hagoort et al., 1993; Neville et al., 1991; Osterhout & Holcomb, 1992). Further exploration of this paradigm may yield a paradigm in which semantic and syntactic processes could be distinguished experimentally.

The current studies use the cross-modal Integration Paradigm introduced in Boland (1991, see also Boland, 1993)[3] as a partial solution to the methodological problem of distinguishing between syntactic and semantic representations. Auditory presentation of a sentence or sentence fragment is followed by a visual target that may or may not be a good continuation of the sentence. The assumption is that as we hear or read a sentence, we immediately begin constructing syntactic and semantic representations of it. Responses to the target word will be faster when the target is consistent with the relevant representation than when it is not. For reasons that are discussed at length in Boland (1991), the ability to integrate the target into the syntactic representation is most relevant when the task is naming, but both the syntactic and semantic representations play a role when the task is lexical decision. Note that naming and lexical decision are

[3]Experiment 1 of the current paper is Experiment 3 in Boland (1991). This experiment was also briefly summarized as Experiment 2 in Boland (1993).

used here to investigate **post-access** effects -- namely, the ease with which the target can be integrated into the sentence context.

The Integration Paradigm contrasts naming and lexical decision tasks based on the evidence that naming is most sensitive to syntactic representations and lexical decision is sensitive to both syntactic and semantic representations. However, this distinction is probably not absolute. For example, recent studies in my own laboratory using only visual representation have not obtained clear patterns of results. The conditions under which this task difference can be obtained must be better understood in order for this paradigm to be fully convincing. The current experiments handle this potential difficulty by including control conditions designed to determine which level(s) of representation are being tapped.

Experiment 1

Experiment 1 was designed to investigate how verb-specific syntactic and thematic knowledge become available to the sentence processing system and how this lexical knowledge influences on-line construction of syntactic and semantic representations of sentences. The focus is on verbs such as *toss*, that have multiple argument structures (i.e., *toss the ball to the child* vs *toss the salad*). If the entire lexical entry is made available at once, all the syntactic and semantic knowledge associated with each of a verb's senses and argument structures would presumably become available. How might the sentence processing system sort out this information and make use of it?

All models of sentence processing assume that at least some aspects of lexical knowledge are used in the initial stages of sentence processing. For example, the Garden Path model assumes that major grammatical category is used (along with phrase structure rules) to assign each incoming word to the phrase structure tree. Interactive models maintain that subcategory and thematic role information are used for initial parsing decisions. In either case, the output of the word recognition system provides crucial input to the sentence processing system. However, the output of the word recognition is often ambiguous and it is not clear how the sentence processing system would handle lexically ambiguous input.

Consider the way lexical access is believed to occur. There is general agreement (Forster, 1979; Marslen-Wilson, 1987) that lexical items are accessed in a bottom-up fashion when linguistic input is perceived; contextual information cannot access lexical items independently, nor does context restrict lexical items from being accessed by the input. Thus, when physical input is ambiguous, multiple lexical forms are accessed, although not always simultaneously (see for example, Van Petten & Kutas, 1987). Context is then used to select the appropriate candidate.[4] However, the most common current views of sentence processing maintain that the parser constructs syntactic representations in serial

[4]However, if one sense of an ambiguous word is accessed earlier than the other senses and it can be quickly integrated into the context, the alternative senses may not be accessed (Rayner & Morris, 1991). For example, Tabossi (1988) found that a subordinate meaning will not be accessed if the appropriate semantic features of the more frequent meaning are primed by the context.

(e.g., Frazier, 1987). A problem arises because a large number of English words are ambiguous at some level. Even if the parser only makes use of syntactic category, what can it do with noun/verb ambiguities like *ring* or adjective/noun ambiguities like *green*? And if the parser makes use of subcategory information in addition to major syntactic category, the ambiguity is multiplied because many verbs allow multiple subcategorization frames. Is the appropriate subcategorization and thematic frame selected in the same way that the appropriate meaning of an ambiguous word is selected?

Experiment 1A uses cross-modal naming and was designed to determine whether multiple subcategorization frames are made available when verbs with multiple senses are recognized. ("Sense" is loosely defined here as a difference in the type of event denoted by the verb.) Experiment 1B uses cross-modal lexical decision to ask the same question about thematic frames. The critical sentences contain verbs with senses that have different numbers of arguments associated with them. For example, the sense of *toss* associated with salads has just two arguments, a subject and an object. In contrast the "throw" sense can have three arguments: subject, direct object, and indirect object. Thus, (1a) is unacceptable[5], but (1b) is fine. Remember, according to the lexical access literature, semantic associates of both senses would be facilitated at the offset of *toss* in both (1a) and (1b). Several hundred milliseconds later, only the contextually appropriate probe would be facilitated. This is because words are accessed in bottom-up fashion -- then context is used to select the most likely sense.

1a) ?**Which salad did Jenny toss Bill?*
b) *Which baseball did Jenny toss Bill?*

Presumably the entire lexical entry for each sense of an ambiguous word is activated, not merely multiple meanings. It ought to be possible to design an experiment that tests for multiple activation of lexical argument structures that is exactly analogous to the semantic priming experiments that test for multiple activation of meanings. If multiple argument structures are made available, then multiple syntactic and thematic representations might initially be formed at words like *toss*. Thus, integrating the target, *BILL*, into the contextually inappropriate representation ought to be equivalent to integrating the target into the contextually appropriate representation, but only if the target is presented during the window of time when both representations are available. Remember, integration effects in naming are likely to reflect syntactic integration, whereas integration effects in lexical decision may reflect both syntactic and semantic integration.

Two control conditions using "unambiguous" verbs are necessary to ascertain whether the task is tapping syntactic representations, semantic representations, or both. The first, illustrated in (2a), used simple transitive verbs that only subcategorize for two arguments: a subject and a direct object. This condition

[5]Note that (1a) is somewhat implausible, but acceptable if the "throw" sense is adopted. Some speakers may also find (1a) acceptable as the short form of *Which salad did Jenny toss for Bill?* There was some variation among the materials regarding how strongly the indirect object biased one meaning over the other and whether or not a benefactive reading was possible.

served as the baseline at which neither syntactic nor thematic integration is possible. The other control condition (2b) distinguishes between semantic integration and syntactic integration by using non-alternating datives. These verbs have three thematic roles and subcategorize for a prepositional indirect object, so that *SAM* is thematically congruent, but syntactically incongruent. If naming latencies for the two control conditions are equivalent and longer than naming latencies for the "baseball" condition, it will verify that naming is only sensitive to syntactic integration, and not thematic integration. Lexical decision latencies will be shorter for the thematically congruent condition if the task is sensitive to thematic integration.

2a) *Which necklace did Nancy touch.. SAM*
b) *Which necklace did Nancy describe.. SAM*

Note that *toss* is an alternating dative that can take a noun phrase indirect object, so the ambiguous three argument condition (illustrated by (1b)) should allow both syntactic and thematic integration under any account. If it is significantly faster than the unambiguous two argument condition (2a), that will provide evidence for an integration effect in either task. In addition, it will provide a standard against which to compare the ambiguous two argument and the unambiguous three argument conditions. These are the two most interesting conditions. If argument structure information follows the same pattern of "activation" as semantic associates, then multiple sets of argument structure information should initially be available. This would be reflected by response times in the ambiguous two argument condition equivalent to those in the ambiguous three argument condition.

Note that it is possible to continue the "incongruent" conditions in such a way that the context+target is a legal string. Some examples are given in (3), below. However, in each of these cases, the target word is the direct object and the wh-phrase is part of an adjunct phrase. A number of researchers have used a variety of paradigms to examine how fronted wh-phrases are analyzed in sentences like those used in this experiment. All the evidence demonstrates that the wh-phrases are assigned the direct object role when a transitive verb is encountered (Clifton et al., 1984; Frazier & Clifton, 1989; Garnsey et al., 1989; Kurtzman, 1989). It is certainly possible that subjects would construct just such a structure when they encounter the target. But doing so would require some reanalysis (of the wh-phrase) and thus response times should be longer in these conditions compared to "congruent" conditions in which no reanalysis is necessary.

3a) *Which salad did Jenny toss Bill the croutons for?*
b) *Which necklace did Nancy touch Sam with?*
c) *Which necklace did Nancy describe Sam wearing?*

The choice of proper names as targets has two consequences. First it insures that all targets are equally unpredictable so that subjects cannot generate potential targets from the context. Second, it required that the traditional lexical decision instructions be modified slightly because the targets were names rather than words. I have assumed that common first names are represented in the lexicon somewhat like nouns -- thus, it is possible to access the lexical item along with the semantic features of human, animate, male, etc.

Experiment 1A

The Concurrent model differs from the Garden Path model and serial Interactive models with regards to the number of syntactic representations that are constructed. The Concurrent model predicts that all subcategorized structures are pursued in parallel. In contrast, the Garden Path model maintains that only the syntactically simplest structure is initially constructed -- without regards to subcategory information. At the other extreme, serial Interactive models maintain that only the most contextually plausible syntactic representation will be constructed.

Experiment 1A tests one hypothesis of the Concurrent model, namely, that syntactic representations corresponding to each subcategorization structure are initially constructed when verbs with multiple subcategorization frames are encountered. This predicts that naming times in the two ambiguous verb conditions will be equivalent, and faster than the unambiguous verb conditions, which are both syntactically incongruent. In contrast, serial Interactive models predict that only the contextually appropriate structure will be constructed, so responses should be faster in the ambiguous three argument condition than the ambiguous two argument condition. These predictions are summarized in (4a) and (4b), respectively.

The predictions of the Garden Path model are less clear. Subcategory information, specifying that a third argument is possible, would not be available at the point when the target is presented. Thus, there might be no representation into which the target could be easily integrated, causing the ambiguous conditions to be equivalent to the unambiguous conditions. Alternatively, if subjects took the target to be part of the sentence, the simplest attachment uses the double object structure associated with (5b), below. Because subcategory information is not available, this structure would be used for both ambiguous and unambiguous verbs -- again predicting no differences across the four conditions as shown in (4c). To insure that the task is tapping the earliest point in processing, when (according to the Garden Path model) subcategory information is not available, the visual target was presented just before the auditory offset of the verb. Note, however, that if subcategory information became available in time to influence the naming response, one might find exactly the pattern predicted by the Concurrent model.

4a) Predictions of the Concurrent model: 3A = 2A < 3U = 2U
b) Predictions of serial Interactive models: 3A < 2A = 3U= 2U
c) Predictions of the Garden Path model: 3A = 2A = 3U = 2U

The unambiguous verb conditions also test the methodological hypothesis that naming is insensitive to semantic congruity. The prediction is that naming times for the two unambiguous conditions will be equivalent, with no advantage for the three argument condition. Furthermore, a difference between the ambiguous and unambiguous argument conditions will provide evidence that subcategorization information is available early. The Concurrent model and the Interactive model both require that subcategory information be available early, while the Garden Path model requires that there be some time-point after word recognition when subcategorization information is not yet available.

Method.

Subjects. Forty undergraduates at the University of Rochester completed the experiment in partial fulfillment of course requirements or for a nominal fee. All were native speakers of English.

Materials. The Ambiguous conditions use ten alternating dative verbs that were judged to have another sense in which they were two argument transitives. For each verb, two versions of a sentence fragment were constructed. The two versions were identical except for the fronted, wh-phrase that was a filler for the direct object gap. In each pair, one of the wh-phrases strongly biased the two argument meaning, and the other strongly biased the three argument meaning. Ten additional sentence fragments were constructed, using the same structure, for the control conditions. Each of these fragments was also made into two versions, which were identical except for the verb. The verb was an unambiguous two argument verb in one version and a non-alternating dative (three argument verb) in the other. A sample set of experimental sentences is shown in (5). The full set is listed in the Appendix. Sentence completion norms were collected on all the experimental contexts to insure that the contexts biased the ambiguous verbs as expected and to insure that my judgments regarding verb subcategorization were appropriate.

5a) Ambiguous 2-Argument: *Which salad did Jenny toss.. BILL*
b) Ambiguous 3-Argument: *Which baseball did Jenny toss.. BILL*
c) Unambiguous 2-Argument: *Which necklace did Nancy touch.. SAM*
d) Unambiguous 3-Argument: *Which necklace did Nancy describe.. SAM*

Ten sentence fragments with ambiguous verbs and ten with unambiguous verbs appeared on each of two lists and the two and three argument conditions of each were rotated between lists. In addition, 58 distractor fragments were constructed, about 30% of which were obviously cut off in mid-sentence. Targets for all trials were common first names, 2-5 letters in length. Targets for the critical trials were all single syllable names, 3-4 letters in length. Naming norms were collected on the targets without contexts. Targets in the ambiguous verb group averaged 393 milliseconds and targets in the unambiguous verb group averaged 390 milliseconds (N=8). There was no difference between the two groups of targets.

The sentence fragments were read into a tape recorder. An attempt was made to read all the critical fragments, as well as those distractors which ended mid-sentence, with neutral (as opposed to sentence-final) intonation. The materials were then digitized using the MacRecorder system. Five millisecond tones at 1000 hz were placed on the non-voice channel approximately 150 milliseconds before the offset of the verb or at the onset of the final consonant. In all cases, the tone occurred after the subjective recognition point of the word.

Procedure. Subjects wore headphones and were seated in front of a computer screen, response box, and microphone. Contexts were presented through the headphones to both ears, and the target names were centered in all capitol letters on the computer screen. Subjects responded by pronouncing the name into the microphone, which was connected to a noise-sensitive switch on the response box. Reaction times were collected from the time when the target came onto the screen until the noise-sensitive switch was triggered. If no response was registered within 2 seconds, the response was considered a time-out. In this paradigm, time-outs usually reflect mechanical trigger-failures rather than slow responses. Yes/No comprehension questions were presented visually after 25% of the trials to insure that subjects attended to the auditory contexts. In no case was it necessary to integrate the target with the context to answer the comprehension question. Subjects completed 10 practice trials, half of which had comprehension questions, before going on to the 78 experimental trials.

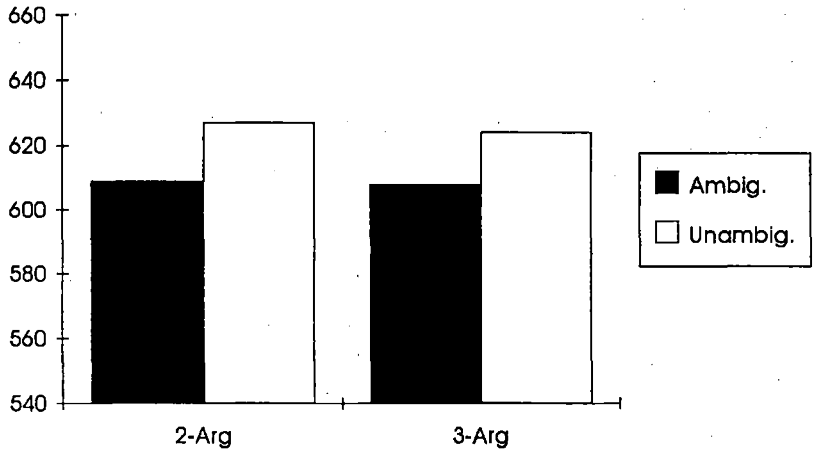

Figure 1. The mean naming latencies for each condition of Experiment 1A are given in milliseconds.

Results.

For each of the critical trials, the onset latency of the naming response was recorded. Time-outs accounted for less than 3% of the data. Within each condition, mean response times were computed by subject and by item. Responses were considered outliers if they were more than 2.5 standard deviations from a subjects mean response time. Outliers were replaced with the boundary value. About 4% of the data were replaced in this way. Mean response times are displayed in Figure 1. Subject and item means were each subjected to a 2(list) x 2(verb type) x 2(argument number) Analysis of Variance (ANOVA). There was a main effect of verb type by subjects and by items [$F1(1,38)=9.58$, $F2(1,16)=8.68$, $p < .01$], with responses to the unambiguous conditions slower than the ambiguous conditions. Importantly, there was no effect of argument number, which would have reflected semantic integration [$Fs < 1.0$], and no interaction between verb type and argument number [$Fs < .10$]. In a planned comparison of the three argument conditions, the ambiguous condition was faster than the unambiguous condition [$F1(1,38)=4.80$, $F2(1,16)=7.84$, $p < .05$]. The difference between the two argument conditions was marginally significant [$F1(1,38)=3.90$, $F2(1,16)=3.63$, $p < .10$]. The two ambiguous conditions did not differ from one another, nor did the two unambiguous conditions.

Discussion.

Both theoretical and methodological predictions were supported. There was no difference between the two unambiguous conditions, and the ambiguous three argument condition was faster than the unambiguous three argument condition. This provides evidence that the task is insensitive to semantic congruity, but sensitive to subcategorization information. The Concurrent model's prediction that both subcategorization frames would be constructed was also supported. Naming times for the ambiguous verb conditions were fast compared to the unambiguous verb conditions. This suggests that the two argument and three argument conditions were both syntactically congruent -- and the two argument condition could not have been syntactically congruent unless the inappropriate subcategorization frame was available. Thus, the pattern of results is inconsistent with the predictions of the Interactive model. Furthermore, the results are not consistent with the Garden Path model unless it is the case that subcategory information becomes available in time to influence the naming response. What makes this unlikely, is the early point at which the target was presented. If subcategorization information is not available in time for the initial parse, it is not clear how it could be available soon enough to influence the naming response. Further evidence against the Garden Path model is provided by Experiment 1B and Experiment 2.

Experiment 1B

The evidence from Experiment 1A, using naming, suggests that the parser constructed a structural representation corresponding to each subcategorization frame of the ambiguous verbs. Experiment 1B, which uses lexical decision, provides an opportunity to replicate that effect (because the lexical decision task is sensitive to syntactic congruity) as well as to test the hypothesis that only a single thematic frame is pursued (because the lexical decision task is also sensitive to semantic congruity). I am assuming that all thematic frames are initially made available based on the evidence that all meanings of ambiguous words are initially made available. However, the Concurrent model predicts that only the thematic frame that is most consistent with the context will be pursued.

The Concurrent model predicts that decision times in the ambiguous verb conditions should be faster than those for the unambiguous verb conditions, as in Experiment 1A. This is the syntactic integration effect. In addition, decision times should reflect semantic integration: the three argument ambiguous condition should be faster than the two argument ambiguous condition because the target is consistent with the context only in the three argument condition. Likewise, decision times for the three argument unambiguous condition should be faster than those for the two argument unambiguous condition because only the former provides a thematic role for the target. These predictions are summarized in (6a).

In contrast, the most straightforward interpretations of the Interactive and Garden Path models predict the same patterns of effects that the models predicted with the naming task, though for different reasons. Interactive models maintain that syntactic and semantic processors work together to construct a single representation, which both naming and lexical decision would presumably tap. The Garden Path model, in contrast, predicts that the simplest syntactic representation is constructed first (ignoring subcategory information), and it is this initial representation that is presumably being tapped. Thus, there should be no difference between the four conditions (as shown in (6c)) because subcategory information is not yet available. Alternatively, if the task taps a later stage of processing, and subcategory information is available in time to influence the response, one of the two syntactic congruity effects illustrated in (7) should be obtained. If the decision task taps a very late stage in processing, and semantic analysis has also occurred, then the pattern predicted by the serial Interactive model (6b) should be obtained.

6a) Predictions of the Concurrent model: 3A < 3U = 2A < 2U
b) Predictions of serial Interactive models: 3A < 3U = 2A = 2U
c) Predictions of the Garden Path model: 3A = 3U = 2A = 2U

7. Alternative Predictions of the Garden Path model:
a) 3A < 3U = 2A = 2U
b) 3A = 2A > 3U = 2U

Method.

Subjects. Forty undergraduates at the University of Rochester completed the experiment in partial fulfillment of course requirements, or for a nominal fee. All were native speakers of English.

Materials. The auditory contexts used in Experiment 1A were used again here. The target list was modified by generating 24 pronounceable, non-names for the distractor trials. (Non-names were non-words and were not homophonous with any common name or word.) Overall, approximately 30% of the trials were non-name trials. Decision norms were collected on the critical targets without any contexts. The targets used with ambiguous verbs averaged 521 milliseconds and the targets used for unambiguous verbs averaged 518 milliseconds (N=8). There were no differences between the two groups.

Procedure. The procedure used in Experiment 1A was used again here, except that subjects were told that a string of letters would appear on the screen and they should decide whether or not it was a real name as quickly as possible. Decisions and their latencies were recorded on a button box labeled with "yes" and "no." If the subject did not respond within three seconds, the program registered a "timeout" and the experiment continued.

Results.

For each of the critical trials, the lexical decision latency was recorded, along with which button was pressed. Subjects made errors on less than 2% of the critical trials. There were no "time-outs" on the critical trials. Within each condition, correct ("yes") mean response times were computed by subject and by item. Mean response times are displayed in Figure 2, below. Outliers were replaced at 2.5 standard deviations as in Experiment 1A. Approximately 3% of the data were replaced in this way.

Subject and item means were subjected to a 2(list) x 2(verb type) x 2(argument number) ANOVA. The data pattern is strikingly different from that obtained using the pronunciation task. The verb type effect, reflecting syntactic congruity, is still observed, but only in the subject analysis: the ambiguous verb conditions are significantly faster than the unambiguous verb conditions by subjects [$F1(1,38)=28.37$, $p < .01$], but not by items [$F2(1,16)=2.25$, $p > .10$]. The planned comparisons of the two argument ambiguous and unambiguous conditions [$F1(1,38)=10.82$, $p < .01$; $F2(1,16)=2.07$, $p > .10$] and three argument ambiguous and unambiguous conditions [$F1(1,38)=4.85$, $p < .05$; $F2(1,16)=1.49$, $p > .10$] were also significant by subjects but marginal by items.

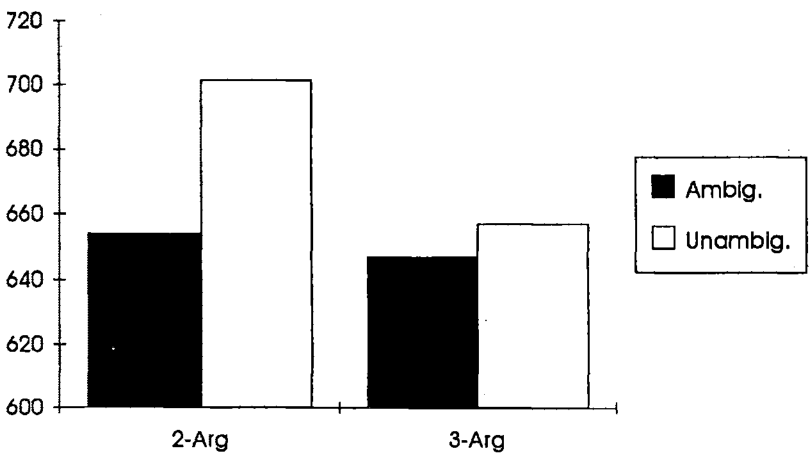

Figure 2. The mean decision latencies for each condition of Experiment 1B are given in milliseconds.

What is striking is the main effect of argument number, with two argument conditions slower than three argument conditions [$F1(1,38)=7.44$, $F2(1,16)=4.69$, $p < .05$]. This reflects semantic integration because there was a thematic role for the target in the three argument condition, but not the two argument condition. Although the effect of argument number did not interact with verb type [$Fs < 1.0$], it appears numerically larger for the unambiguous verbs. In a planned comparison of the unambiguous verbs, this difference was reliable by subjects [$F1(1,38)=5.21$, $p < .05$] and marginally reliable by items [$F2(1,16)=4.19$, $p < .10$]. However, the difference was not reliable for the ambiguous verbs [$Fs < .10$].

Discussion.

The pattern of results confirmed that the decision task is sensitive to both syntactic and semantic representations, and that thematic and contextual information are used very early to guide the semantic representation. Evidence that the task is sensitive to thematic information is provided by the unambiguous control conditions. The unambiguous three argument condition was faster then the unambiguous two argument condition because the three argument condition offered a thematic role for the target. The main effect of verb type, seen previously in Experiment 1A, is again evidence of syntactic integration. Responses to the unambiguous verb conditions were comparatively slow because they are not syntactically congruent. By contrast, the ambiguous three argument condition is syntactically congruent on all accounts, and the ambiguous two argument condition is syntactically congruent if the alternative subcategorization frame is available.

The main effect of argument number indicates that only the contextually appropriate thematic frame was pursued. Thus, it was easier to integrate the target in the three argument conditions compared to the two argument conditions. Although the size of the effect appears larger for the unambiguous verbs than the ambiguous verbs, the effect of argument number did not interact with verb type.

The results can best be interpreted as follows: multiple subcategorization frames are pursued without regards to context, but the availability of thematic frames is regulated by (top-down) contextual information.

Summary of Experiment 1

Experiment 1 provides further evidence that cross-modal naming and lexical decision are differentially sensitive to syntactic and semantic representations. Furthermore, the experiment provides new data on the relationship between syntactic and semantic processing, and how verb argument structure is used by the sentence processing system. The results suggest that verb argument structure is accessed in parallel, much like multiple meanings of ambiguous words. However, the syntactic and semantic systems deal with this parallel information differently. The syntactic system automatically constructs a representation consistent with each subcategorization frame, but the semantic system uses context to select the most likely thematic frame to pursue. In such a system, garden paths would occur only when the thematic system pursued the incorrect interpretation (because the context was misleading or uninformative). In this case, the alternative syntactic frames might be used to identify an alternative analysis.

An alternative line of explanation for the data in Experiment 1 must also be considered. Suppose that the two argument subcategorization and thematic frames of the ambiguous verbs were ruled out, brute force fashion, by the target word. The two argument ambiguous condition would be syntactically congruent, but implausible, and the three argument ambiguous condition would be syntactically congruent and plausible. Because both would be syntactically congruent there would be no difference in naming times, but the plausibility difference would be reflected in the lexical decision times. This explanation is difficult to rule out, but it is unlikely because there is no reason to think that subjects were forcibly trying to integrate the target word with the contexts. It was assumed that the target would be integrated automatically only if it was congruent. Fully half of the experimental contexts were complete sentences without integrating any probe word so it is unclear why subjects would adopt a strategy of forcibly integrating the probe word. Further evidence against this explanation is provided by Experiment 2. Experiment 2 is a naming experiment that uses the same materials, but probes at time points after the offset of the verb. At late time points, the ambiguous two argument condition is no longer as fast as the ambiguous three argument condition. Thus, it is clear that at 150 and 300 milliseconds post offset, bottom up evidence of the three argument structure does not force that analysis. It is unlikely then, that such a process occur at earlier probe times.

Experiment 2

Experiment 1 suggests that all subcategorized structures are automatically constructed in parallel, but a single thematic frame is selected using contextual information. Experiment 2 was designed to explore the relationship between syntactic and semantic representations over time by testing the availability of the alternative syntactic frame at different time points. This was accomplished using the naming task in a cross-modal, multi-ISI (inter-stimulus-interval) design. The temporal relationship between the offset of the auditory context and the appearance of the next target was varied from 150 milliseconds prior to offset to 300 milliseconds post offset.

I have assumed that the naming data and the lexical decision data reflect the same time point in processing. Thus, the results of Experiment 1 demonstrate that a single interpretation has been selected **before** the syntactic ambiguity is resolved. However, it is also possible that the naming task captures the sentence processor at an earlier point than does lexical decision -- that, in fact, the semantic processor cannot develop an interpretation until a single syntactic structure is passed up by the syntactic processor. This possibility calls into question the task difference, itself, because it follows that if the naming response was slowed down the task would be sensitive to semantic integration. The latter possibility is ruled out by the naming data from the current experiment. We will see that even 450 milliseconds later (300 milliseconds post-offset), there is no evidence of a semantic effect with the naming task.

This leaves the question of how the subcategorization ambiguity is resolved. According to the Concurrent model, once syntactic and semantic representations are developed, they are compared, and inconsistent representations are discarded. This leads to the prediction that argument number and ISI should interact when the verb is ambiguous. Specifically, the ambiguous two argument condition should be fast at short ISI's and slow at long ISI's with the naming task. (It should be slow at all ISI's using lexical decision.) On the other hand, if syntactic parallelism is maintained until there is bottom-up evidence for one subcategorization over the other, one would expect the ambiguous two argument condition to remain fast in the naming task, because integration of the probe word should constitute bottom-up evidence for the three argument subcategorization.

The data pattern for the ambiguous verbs that is predicted by the Concurrent model is identical to the pattern predicted if naming is sensitive to semantic congruity at long ISI's. Therefore, the control conditions are again crucial. If naming is truly sensitive only to syntactic congruity, argument number should not interact with ISI when the verb is unambiguous. That is, the unambiguous three argument condition should not be faster than the unambiguous two argument condition at long ISI's. However, if the source of the task difference between naming and lexical decision is in the relative timing of the response -- and naming was tapping an earlier representation than lexical decision in Experiments 1 and 2 -- then the pattern of naming responses should mimic the lexical decision task at long ISI's, and the unambiguous three argument condition should be faster than the unambiguous two argument condition.

Method.

Subjects. Eighty undergraduate students from the Ohio State University served as subjects, 20 in each of the ISI conditions. All were naive to the experimental hypothesis and were native speakers of English. Subjects participated to fulfill part of their course requirement in introductory psychology.

Materials. The materials from Experiment 1 are used again here, although they were spoken by a different person, re-randomized and assigned different targets.

Procedure. The auditory sentence contexts were digitized and edited as in Experiment 1. As before, sync tones were set at approximately 150 milliseconds before the offset of the last word in the auditory context. However, the temporal relationship between the sync tone and the presentation of the target word was manipulated between subjects so that there were 4 ISI conditions. The target was presented at the sync tone for the "-150 ISI" group, 150 milliseconds after the sync tone for the "0 ISI" group, 300 milliseconds after the sync tone for the "150 ISI" group, and 450 milliseconds after the sync tone for the "300 ISI" group.

The only other procedural change was that the fastest and most accurate subject in each task was awarded a $10 prize. This incentive produced somewhat faster response times than were seen in Experiment 1.

Results.

The results are summarized in Figure 3. Outliers were replaced at 2.5 standard deviations and subject and item means were computed as in the previous experiments. These means were first subjected to a 2(list/item group) x 2 (verb type) x 2(argument number) x 4(ISI) analysis of variance. In the subject analysis, list and ISI were between factors and verb type and argument number were within factors. In the item analysis, item group and verb type were between factors, while argument number and ISI were within.

The predicted effect of verb type (with the ambiguous verb conditions faster than the unambiguous verb conditions) was obtained in the subject analysis [F1(1,72)=13.04, p < .001], but was not reliable in the item analysis [F2(1,16)= 1.93, p > .10]. Importantly, this effect did not interact with ISI [F1 & F2 < 1.0], demonstrating that the syntactic congruity effect was maintained across time. In addition, there was no main effect of argument number [F1 & F2 < .10], which would have reflected a semantic congruity effect. The main effect of ISI was reliable in the item analysis, but not in the subject analysis [F1(3,72)=1.39, p > .10; F2(3,14)=16.14, p < .01].

Unfortunately, the predicted three-way interaction between verb type, argument number, and ISI was not obtained [F1 & F2 < 1.0]. This interaction was predicted because argument number and ISI should interact for ambiguous verbs, but not for unambiguous verbs. Although the predicted pattern was obtained at the last three probe positions, with the ambiguous two argument condition gradually becoming less available, the ambiguous two argument condition was inexplicably (and non-reliably) slower than the ambiguous three argument condition at the first probe position. Thus, instead of the predicted three-way interaction, only a two-way interaction between verb type and argument number was obtained [F1(1, 72)=7.08, p < .01; F2(1,16) = 6.32, p < .05].

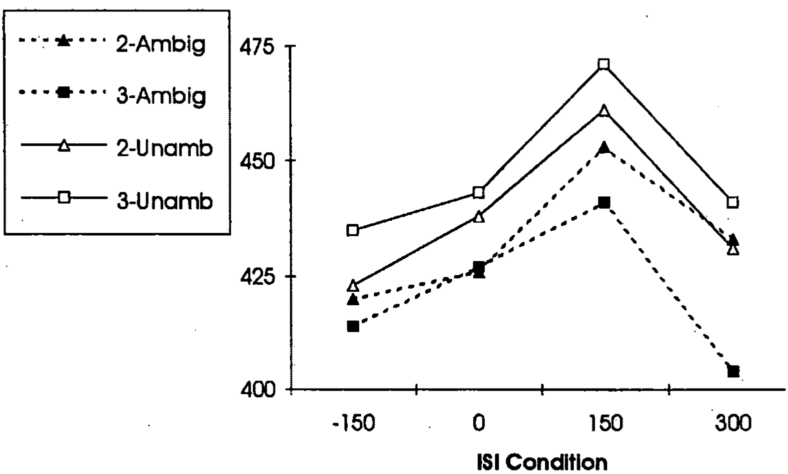

Figure 3. The mean naming latencies for each condition of Experiment 2 are given in milliseconds.

Because the model makes predictions about differences between ISI conditions, separate 2(list) x 2(verb type) x 2(argument number) ANOVAs were done on each ISI group. At the earliest ISI (-150), the data was quite noisy, and no main effects or interactions reached significance. By the offset of the verb (0 ISI), the main effect of verb type was reliable by subjects [F1(1,18)=5.44, p = .10], but not by items [F2(1,16)=1.48, p > .10]. Shortly after verb offset, at 150

ISI, the effect of verb type was only marginally reliable by subjects [$F1(1,18) = 3.99$, $p < .10$; $F2(1,16)=1.53$, $p > .10$], and it failed to reach significance completely by 300 milliseconds post offset [$F1(1,18)=2.48$, $F2(1,16)=2.26$, $p > .10$]. At the last ISI, there was also an interaction between verb type and argument number [$F1(1,18)=7.18$, $F2(1,16)=5.83$, $p < .05$]. There was never an effect of argument number [$Fs < 1.0$].

Planned comparisons of the ambiguous and unambiguous three argument conditions demonstrated that the ambiguous three argument condition was reliably faster by subjects and marginally faster by items at 150 milliseconds post-offset ($\alpha = .05$). The difference was reliable in both subject and item analyses at 300 milliseconds post offset. The unambiguous three argument condition never differed reliably from the unambiguous two argument condition in either the subject or the item analyses. The only ISI at which the ambiguous three argument condition was reliably faster than the ambiguous two argument condition was the 300 ISI condition.

Discussion.

Experiment 2 used the naming task to examine the state of the syntactic representation(s) at various points in time. There were two crucial predictions, one theoretical and one methodological. First, if the syntactic and semantic representations are compared and inconsistent analyses discarded, only the contextually appropriate syntactic frame should be available at long ISI's, and the ambiguous two argument condition should be slow. This is the prediction of the Concurrent model. In contrast, if syntactic parallelism is maintained until there is bottom-up evidence for one subcategorization frame over the others, the ambiguous two argument condition should remain fast because integration of the probe word would constitute bottom-up evidence for the three argument subcategorization. The pattern predicted by the Concurrent model was obtained over the last three probe positions. Note that this pattern does not constitute evidence that naming was sensitive to semantic congruity at later time points. Although that is one possible interpretation of the data for the ambiguous conditions, it is ruled out by the data for the unambiguous conditions.

The second crucial prediction was that the unambiguous three argument condition would never differ from the unambiguous two argument condition. Recall that in Experiment 1, they did not differ with the naming task, but the three argument condition was faster with lexical decision. The effect in lexical decision was attributed to the task's sensitivity to semantic congruity, because the unambiguous three argument verbs allowed a third thematic role that was consistent with the semantic features of the target. However, there is an alternative explanation consistent with the Garden Path model that must be ruled out. The alternative is that lexical decision task taps processing at a later stage than does naming. Under this account, both naming and lexical decision are sensitive to semantic congruity in principle, and semantic effects would be seen in naming if there were enough time for the semantic information to become available. To rule out this explanation, the unambiguous three argument condition must remain slow at all ISI's. As seen in Figure 3, this pattern was obtained; there was no difference between the unambiguous three argument condition and the unambiguous two argument condition, even at the longest ISI. This demonstrates that naming was insensitive to thematic congruity in this paradigm.

Somewhat surprisingly, no clear pattern emerged at the first ISI, and this appeared to disrupt the predicted three-way interaction. Although the -150 ISI was used in Experiment 1, the data pattern found in Experiment 1 is seen here at verb offset. This discrepancy is probably due to the relatively short naming latencies observed here compared to those in Experiment 1. Thus, responses at verb offset are presumed to reflect the same stage of processing observed in Experiment 1. As before, both subcategorization frames for the ambiguous verbs were available. Unlike Experiment 1, there does appear to be a difference between the unambiguous control conditions -- the three argument condition is slower, not only here, but at each ISI. This difference between the control conditions was not predicted by any of the models, and was not reliable at any ISI.

The most interesting data is from the longer ISI conditions. By 300 milliseconds post-offset, only the contextually appropriate subcategorization frame was available, and the predicted interaction between verb type and argument number was obtained. Crucially, there is never an effect of argument number, which would have reflected semantic congruity. The apparently gradual decrease in the availability of the alternative subcategorization frame is consistent with a gradual decay in its activation level once the appropriate syntactic representation has been successfully matched to the semantic representation.

General Discussion

This set of experiments used cross-modal naming and lexical decision in the Integration Paradigm to explore how subcategory and thematic information is used by the sentence processing system. Because this combinatory lexical information must be used by the sentence processing system at some point, investigations into how they are used will also provide insight into the relationship between syntactic and semantic processing. A Concurrent model of sentence processing was proposed, and some of its predictions were tested against the competing interactive and modular models. The evidence provided here supported the Concurrent model.

The Concurrent model of sentence processing maintains that, when a verb is recognized, all of its subcategory and thematic information is accessed. The syntactic processing system uses the subcategory information to construct all the subcategorized structures in parallel, without consulting semantic or contextual sources of guidance. Meanwhile, the semantic processing system uses contextual information, along with any preliminary output from the syntactic system, to select the most likely interpretation. Once semantic and (parallel) syntactic representations are constructed, they are compared to eliminate inconsistent analyses. This is similar to other processing models in which multiple syntactic structures are proposed in parallel (e.g., Crain & Steedman, 1985; Gorrell, 1989 & 1991; Hickok, 1993; MacDonald, 1993; McElree, 1993), except that the Concurrent model makes the explicit claim that a single interpretation is sometimes constructed before the syntactic system has identified the appropriate structure. This claim was supported by the lexical decision results of Experiment 1B, especially when compared against the naming results of Experiment 2. Subcategorization information, thematic information, and contextual information were all used very early, but only subcategorization information influenced the initial syntactic representations. All three sources of information were used to guide the semantic interpretation. This pattern of data is consistent with Parallel Syntax models that allow a single semantic representation to be constructed before a single syntactic analysis is selected.

The evidence from both experiments suggests that, under appropriate experimental conditions, naming is sensitive to syntactic representations and relatively insensitive to semantic representations. In contrast, lexical decision is sensitive to both syntactic and semantic representations. However, this task difference is not as reliable as one would hope. For instance, Gorrell (1991) found that lexical decision was *insensitive* to animacy violations, and Duffy and colleagues (e.g., Duffy et al., 1989) have found that naming was *sensitive* to plausibility. It is important to note, particularly when comparing these experiments to similar studies, that the targets were presented to the subjects without warning, and prior to the offset of the final context word. Many studies have presented a warning signal before the target or offset the target from the context by half a second or more, altering the nature of the task. Furthermore, the paradigm may be limited in utility to cross-modal presentation. Experiments in my own laboratory that have used visual presentation have produced noisy results without clear task differences.

Nonetheless, the Integration Paradigm may prove useful because it has the important capability of mapping changes in representations over time. This feature was exploited in Experiment 2, to examine changes in the syntactic representations. The results suggested that the contextually inappropriate subcategorization frame gradually became less available. I have suggested that the alternate syntactic representation could be used to recover from a garden path, but as yet I have offered no evidence of this. Further research is necessary to determine if, in fact, alternative subcategorization frames do serve as a mechanism for recovery from a garden path by manipulating the point at which disambiguating information becomes available. I predict that recovery would be more efficient if disambiguating material appeared while the corresponding syntactic representation was still available.

References

Bates, E. & MacWhinney, B. (1982) Functionalist approaches to language acquisition. In *Language acquisition: The state of the art* (E. Wanner & L.R. Gleitman, editors) Cambridge: Cambridge University Press.

Boland, J.E. (1991) *The use of lexical information in sentence processing.* Doctoral dissertation. The University of Rochester, Rochester, NY.

Boland, J.E. (1993) The role of verb argument structure in sentence processing: Distinguishing between syntactic and semantic effects. *Journal of Psycholinguistic Research,* 22, 109 - 132.

Clifton, C., Jr., Frazier, L. & Connine, C. (1984) Lexical expectations in sentence comprehension. *Journal of Verbal Learning and Verbal Behavior,* **23**, 696-708.

Crain, S. & Steedman, M. (1985) On not being led up the garden path: The use of context by the psychological syntax processor. In *Natural language parsing: Psychological, computational, and theoretical perspectives* (D.R. Dowty, L. Karttunen, & A.M. Zwicky, editors) Cambridge: Cambridge University Press.

Duffy, S. A., Henderson, J. M., & Morris, R.K. (1989) Semantic facilitation of lexical access during sentence processing. *Journal of Experimental Psychology: Learning, Memory, & Cognition,* **15**, 795-801.

Ferreira, F. & Henderson, J. M. (1990) The use of verb information in syntactic parsing: Evidence from eye movements and word-by-word self-paced reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **16**, 555-568.

Forster, K.I. (1979) Levels of processing and the structure of the language processor. In *Sentence processing: Studies presented to Merrill Garrett* (W.E. Cooper & E.C.T. Walker, editors) Hillsdale, NJ: Erlbaum.

Frazier, L. (1987) Theories of syntactic processing. In *Modularity in Knowledge Representation and Natural Language Processing* (J.L. Garfield, editor) Cambridge, MA: MIT Press.

Frazier, L. (1989) Against lexical generation of syntax. In *Lexical Representation and Process* (W.D. Marslen-Wilson, editor) Cambridge, MA: MIT Press.

Frazier, L. & Clifton, C., Jr. (1989) Successive cyclicity in the grammar and the parser. *Language and Cognitive Processes,* **4**, 93-126.

Garnsey, S.M., Tanenhaus, M.K., & Chapman, R.M. (1989) Evoked potentials and the study of sentence comprehension. *Journal of Psycholinguistic Research,* **18**, 51-60.

Gorrell, P. (1989) Establishing the loci of serial and parallel effects in syntactic processing. *Journal of Psycholinguistic Research,* **18**, 61-74.

Gorrell, P. (1991) Informational encapsulation and syntactic processing. *NELS proceedings,* **21**.

Hagoort, P. Brown, C., & Groothusen, J. (1993) The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes,* **8**, 439-484.

Hickok, G. (1993) Parallel parsing: Evidence from reactivation in garden-path sentences. *Journal of Psycholinguistic Research,* **22**, 239 -250.

Kurtzman, H.S. (1989) Locating Wh-traces. In *The MIT Parsing Volume, 1988-89* (C. Tenney, editor).

Marslen-Wilson, W.D. (1987) Functional parallelism in spoken word-recognition. *Cognition,* **25**, 71-102.

Marslen-Wilson, W.D. & Tyler, L.K. (1987) Against Modularity. In *Modularity in Knowledge Representation and Natural Language Processing* (J.L. Garfield, editor), Cambridge, MA: MIT Press.

MacDonald, M.C. (1993) The interaction of lexical and syntactic ambiguity. *Journal of Memory & Language*, **32**, 692 - 715.

McElree, B. (1993) The locus of lexical preference effects in sentence comprehension: A time-course analysis. *Journal of Memory and Language*, **32**, 536-571.

Mitchell, D.C. (1989) Verb-guidance and other lexical effects in parsing. *Language and Cognitive Processes*, **4**, 123-154.

Neville, H., Nicol, J., Barss, A., Forster, K.I., & Garrett, M.F. (1991) Syntactically based sentence processing classes: Evidence from even-related brain potentials. *Journal of Cognitive Neuroscience*, **3**, 151-165.

Osterhout, L. & Holcomb, P.J. (1992) Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory & Language*, **31**, 785-806.

Rayner, K., & Morris, R.K. (1991) Comprehension processes in reading ambiguous sentences: Reflections from eye movements. In *Understanding word and sentence* (G.B. Simpson, editor), North-Holland: Amsterdam.

Tabossi, P. (1988) Accessing lexical ambiguity in different types of sentential contexts. *Journal of Memory and Language*, **27**, 324-340.

Tanenhaus, M.K., Carlson, G.N., & Trueswell, J.C. (1989) The role of thematic structures in interpretation and parsing. *Language and Cognitive Processes*, **4**, 211-234.

Van Petten, C. & Kutas, M. (1987) Ambiguous words in context: An event-related potential analysis of the time course of meaning activation. *Journal of Memory and Language*, **26**, 188-208.

Waltz, D.L. & Pollack, J.B. (1985) Massively parallel parsing: A strongly interactive model of natural language interpretation. *Cognitive Science*, **9**, 51-74.

Appendix

The critical materials used in the two experiments are listed here. The first version of each item is the two argument-condition and the second is the three argument-condition.

A. <u>Ambiguous Verbs</u>.

1. Which salad/baseball did Jenny toss
2. Which chapter/letter did Howard write
3. Which prison sentence/fancy dessert did Henry serve
4. Which dark alley/salt shaker did Linda pass
5. What dress/fee did Mrs. Smith charge
6. What town/package did Mr. Simpson leave
7. Which victims/seat did Martin save
8. What kind of tantrum/frisbee did Becky throw
9. Which excuse/gift did Robyn buy
10. Which new magazine/bedtime story did Alice read

B. <u>Unambiguous Verbs</u>.

1. Which necklace did Nancy inspect/describe
2. Which poem did Martha finish/dedicate
3. Which friend did Leonard insult/introduce
4. Which quote did Kathy underline/explain
5. Which pie did Mrs. Jones smell/recommend
6. Which package did Cindy open/deliver
7. Which notebook did Patty damage/return
8. Which secret recipe did Nancy follow/entrust
9. Which hotel did Mr. Peterson examine/mention
10. Which task did Larry despise/demonstrate